Quantifying uncertainty in mean earthquake interevent times for a finite sample

Quantifying uncertainty in mean earthquake interevent times for a finite sample

M. Naylor,[1] I. G. Main,[1] and S. Touati[1]

Received 13 June 2008; revised 4 September 2008; accepted 27 October 2008; published 29 January 2009.

[1] Seismic activity is routinely quantified using means in event rate or interevent time. Standard estimates of the error on such mean values implicitly assume that the events used to calculate the mean are independent. However, earthquakes can be triggered by other events and are thus not necessarily independent. As a result, the errors on mean earthquake interevent times do not exhibit Gaussian convergence with increasing sample size according to the central limit theorem. In this paper we investigate how the errors decay with sample size in real earthquake catalogues and how the nature of this convergence varies with the spatial extent of the region under investigation. We demonstrate that the errors in mean interevent times, as a function of sample size, are well estimated by defining an effective sample size, using the autocorrelation function to estimate the number of pieces of independent data that exist in samples of different length. This allows us to accurately project error estimates from finite natural earthquake catalogues into the future and promotes a definition of stability wherein the autocorrelation function is not varying in time. The technique is easy to apply, and we suggest that it is routinely applied to define errors on mean interevent times as part of seismic hazard assessment studies. This is particularly important for studies that utilize small catalogue subsets (fewer than ~1000 events) in time-dependent or high spatial resolution (e.g., for catastrophe modeling) hazard assessment.

Citation: Naylor, M., I. G. Main, and S. Touati (2009), Quantifying uncertainty in mean earthquake interevent times for a finite sample, *J. Geophys. Res.*, *114*, B01316, doi:10.1029/2008JB005870.

1. Introduction

[2] The earthquake record of the digital age is not yet sufficiently well temporally sampled to define a long-term average in recurrence rates, primarily due to the extreme events having recurrence times greater than the digital record era of ~35 years. For example, the Sumatran earthquake was sufficient to modify the best fit statistical model of the global frequency-moment distribution from a gamma distribution before the event to a pure power law Gutenberg-Richter fit after the event [e.g., *Main et al.*, 2008]. Here we explore the effect of such limited temporal sampling on the distribution of interevent times and specifically examine its rate of convergence to a central limit as the sample window is increased.

[3] Statistical convergence and the reduction of errors with increasing sample size are essentially the same phenomenon. The accuracy of any mean estimate derived from a finite sample size N relies on the central limit theorem which tells us that in the limit of infinite sample size, for uncorrelated data, the distribution of sample means will tend to a normal distribution independent of the form of the "parent" distribution [e.g., *Laplace*, 1812; *Bouchaud and Potters*, 2001; *Ross*, 2003]. As a corollary the rate at which the sample mean converges to the parent distribution mean (from an infinite sample) is $1/\sqrt{N}$. In practice, for uncorrelated data, $1/\sqrt{N}$ Gaussian convergence is observed for relatively small sample sizes ($N > 10$) for a wide range of parent distributions provided the moments are finite. However, the sample mean will converge slower than $1/\sqrt{N}$ if correlations exist in the data.

[4] The applicability of such analysis also requires the parent distributions to be stationary. On a timescale of a few Ma, much greater than the repeat times of the largest earthquakes, the far field drivers of plate motion are remarkably stationary [*DeMets*, 1993]. As a consequence, time-independent hazard assessments are made by assuming the long-term behavior is stationary. How the resulting strain is accommodated locally will depend upon the structural setting, for example in collisional mountain belts the continuous driving of plate convergence results in punctuated deformation of individual thrusts within a fold and thrust belt $\sim 10^6 - 10^7$ years [*Naylor and Sinclair*, 2007]. The short-term nature of the catalogues (~35 years of digital data) means that the degree of stationarity is impossible to assess, and introduces significant uncertainty in estimates of long-term hazard.

[5] We demonstrate here how three separate effects limit the statistical convergence of earthquake event rates and interevent times. These are (1) correlations in the form of aftershocks, (2) the skewness of the underlying distribution, and (3) finite sample size. First we demonstrate how the

[1]School of GeoSciences, University of Edinburgh, Edinburgh, UK.

0148-0227/09/2008JB005870$09.00

Table 1. Ranges and Properties of the Earthquake Catalogues Used Here

Catalogue Property	PDE	NZ	SCEC
Date range	12 January 1971–12 January 2005	12 January 1969–12 January 2006	12 January 1968–12 January 2006
Lower magnitude threshold, m_c	5.0	4.0	3.0
Total number of events available, N_{events}	49,196	14,745	12,128
Number of samples of size 4096 that can be drawn from catalogue subset, N_{4096}	12	3	2

statistical properties of mean interevent times are relatively more statistically stable than event rate as a metric for characterizing background activity.

2. Non-Gaussian Convergence of Event Rates and Interevent Times

[6] In this section we discuss the nature of convergence observed in three different earthquake catalogue subsets for New Zealand, Southern California and global PDE (Table 1) that we will be analyzing further in this study. *Main et al.* [2008] demonstrated that the running monthly mean of earthquake event rate is steadily increasing in the global CMT catalogue; we extend this analysis here. Monthly event rates, $\dot{N}_{monthly}$ are shown for each catalogue in Figures 1a–1c; they all confirm that lower event rates are more frequently observed. Global monthly event rates (Figure 1c) do not look like a scaled up version of the

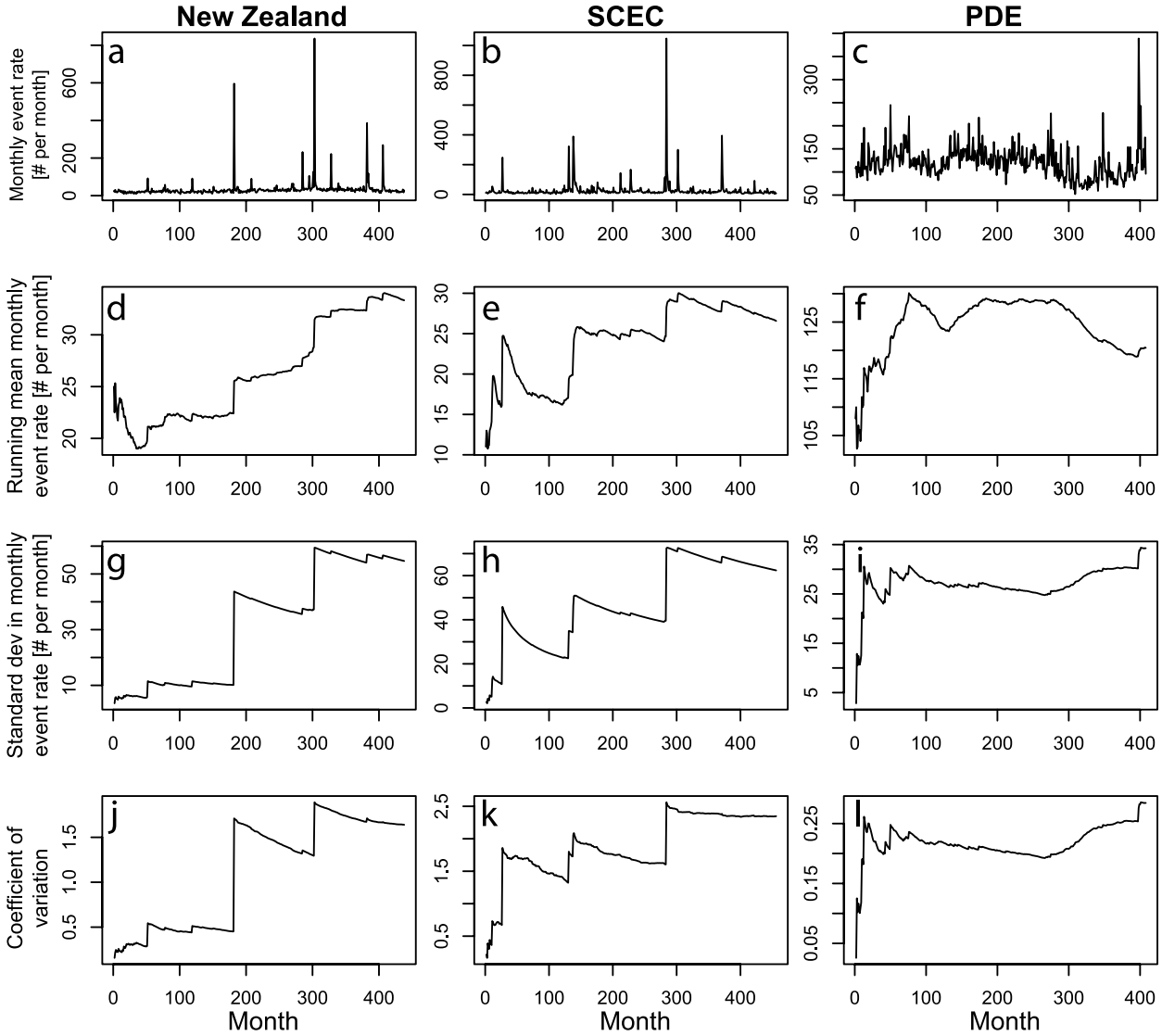

Figure 1. Examples of earthquake data that demonstrate the non-Gaussian convergence of mean event rates for New Zealand, Southern California, and a global catalogue in agreement with the analysis. (a–c) The incremental monthly event rates. (d–f) Steadily rising running mean event rates with the most rapid changes in event rate coincident with intermittent spikes in the monthly event rate. (g–i) Non-Gaussian convergence of event rates characterized by a series of intermittent jumps that correlate with high monthly event rates. (j–l) The coefficient of variation for the running mean monthly event rates.

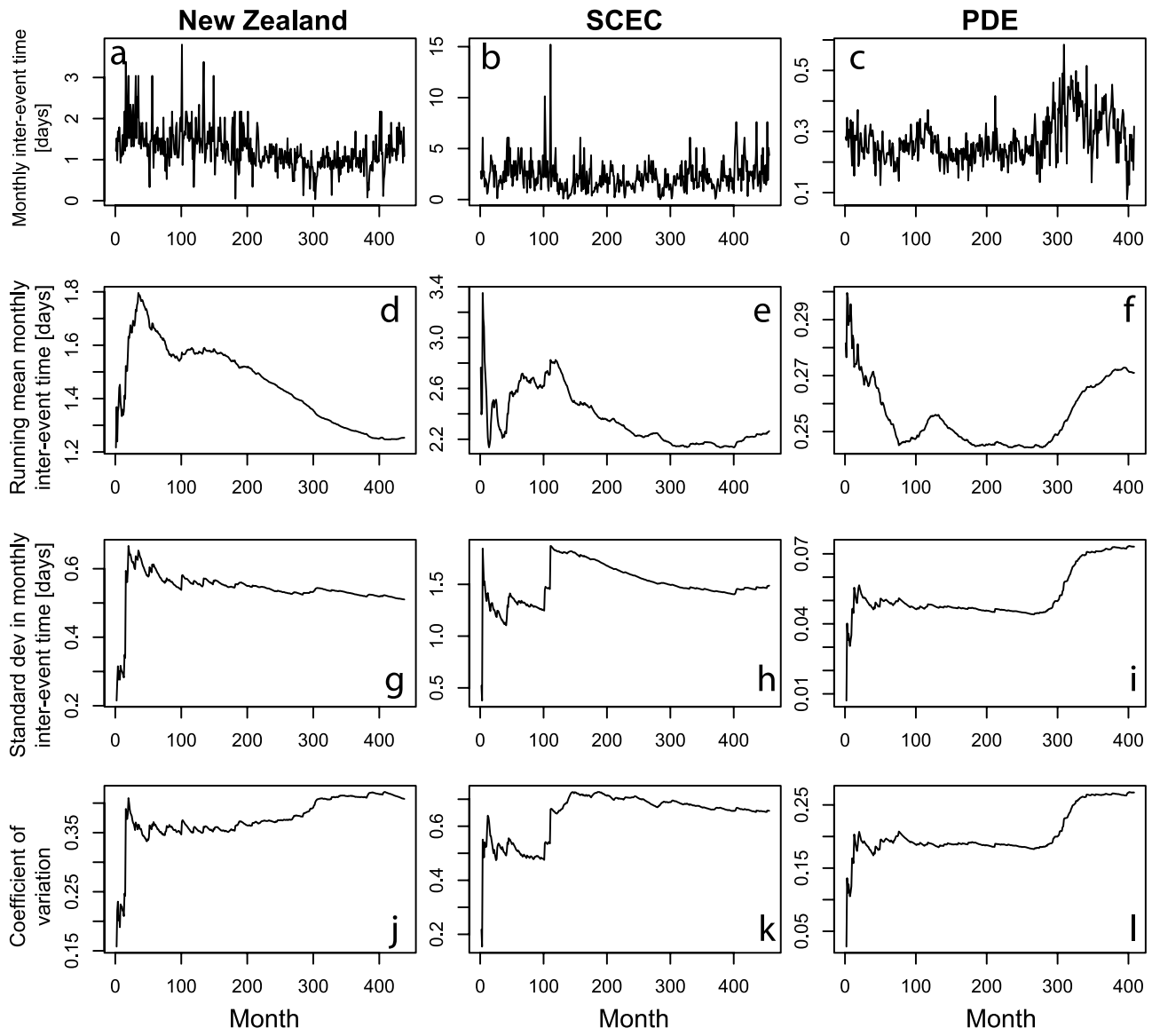

Figure 2. Interevent time. (a–c) Monthly, (d–f) running monthly mean, and (g–i) standard deviation of interevent times making up the running monthly mean. (j–l) Coefficient of variation for the running monthly mean interevent times.

regional rates (Figure 1a and 1b) because triggering dominates in the regional catalogues, where the spatial extent of the analysis region is on the order of the spatial earthquake triggering correlation length (∼100km) [e.g., *Huc and Main*, 2003]. Since high monthly event rates occur less frequently, i.e., there is significant negative skew in the event rate probability density function, lower event rates are sampled more frequently and the running mean monthly event rate will, on average, converge to the mean from below (Figures 1d–1f) in agreement with *Main et al.* [2008]. Jumps in cumulative mean event rate are coincident with the infrequent extreme high monthly event rates. Similarly, the standard deviation of the monthly event rates used to generate the running mean (Figures 1g–1i) increases dramatically when these high event rates are sampled. This effect is so strong for the small spatial sample regions (New Zealand and SCEC) that the standard deviation of event rates becomes greater than the mean (Figures 1j and 1k). In contrast, the global sample has a coefficient of variation less than 1 (Figure 1l). Thus convergence of event rates will occur when such extreme event rates no longer significantly affect either the mean or standard deviation of the event rate. This will only occur once the variance in the tail at high event rates is well sampled; in other words we must wait for the largest events or use a declustering algorithm to remove extreme fluctuations to produce a well-behaved metric. Due to the effect of these extreme events on the running monthly event rate, convergence is punctuated and highly non-Gaussian ($N \gg 10$ for convergence).

[7] An alternative way to view the same data is to convert it to a mean monthly interevent time measured in days $\overline{\delta t}_{monthly} = 30.44/\dot{N}_{monthly}$, where 30.44 is the average monthly duration measured in days (Figures 2a–2c). This transformation changes the weights of the events such that

the frequently occurring low rates now carry the most weight and the extreme events carry the least. The difference in weights between the mean event rate and mean interevent time is related to the difference between arithmetic means and harmonic means. In calculating the running mean the occurrence of an extreme event only increments the sum of past interevent times by a small amount and the number of events by 1 resulting in a small change in the mean. This is in contrast to running mean event rate where an extreme event added a very large number to the sum of past event rates resulting in step changes in the mean. Thus mean interevent times tend to converge from above, rather than below, and also to mitigate the impact of the extreme events on the running mean (Figures 2d–2f). This is reflected in the coefficient of variation which is much more stable and less than 1 for all catalogues (Figures 2j–2l). Thus the uncertainties on earthquake interevent times derived purely from the underlying parent distribution are much lower than for event rates, even though one is completely derived from the other.

[8] In summary, interevent times mitigate the effect of extreme event fluctuations, a desirable property for a metric characterizing background activity. Thus in the remainder of this paper we investigate convergence of earthquake mean interevent times in order to quantify errors on the mean.

3. Theory

[9] In this section we review and develop the theory required to understand convergence of mean earthquake interevent times. First, we define some terminology. A parent distribution is a probability density function from which values of a random variable are drawn, which we assume to be stationary over very long timescales in our application. A collection of random variables drawn from the parent distribution is a sample. The sample mean is the mean of the random variables in a sample. The sample length, N is the number of values used to calculate the sample mean. The distribution of sample means is a histogram of many sample means taken from the parent distribution. Here our time series is a sequence of interevent times. The relation between these definitions and the earthquake problem with which we apply it to in this study is illustrated in Figure 3, where the known parent distribution described above is replaced by the measured earthquake time series.

3.1. Null Hypothesis for Convergence Without Correlations: Random Sampling From the Gamma Distribution

[10] Two key results of probability theory underlie this study.

[11] (1) The ''Strong Law of Large Numbers'' [e.g., *Poisson*, 1837; *Ross*, 2003] gives the intuitive relation that for a sequence of independent random variables $x_1, x_2, \ldots$ with a common stationary parent distribution, and an expectation of the mean $\langle x_i \rangle = \mu$, then with probability 1,

$$\frac{x_1 + x_2 + x_3 + \ldots + x_N}{N} \rightarrow \mu \qquad \text{as} \qquad N \rightarrow \infty. \tag{1}$$

[12] In other words, the mean of a randomly sampled distribution tends to the mean of the parent distribution as $N \rightarrow \infty$; it is this result that allows us to expect that we gain a better estimate of the mean interevent time with a longer earthquake record.

[13] (2) The ''central limit theorem'' [*Laplace*, 1812] gives the less intuitive relation that the distribution of sample means drawn randomly and independently from a parent distribution, provided its first and second moments are finite, tends to a Gaussian distribution as $N \rightarrow \infty$, independent of the parent distribution [e.g., *Ross*, 2003]. This drives Gaussian convergence. Formally, the distribution of

$$\frac{x_1 + x_2 + x_3 + \ldots x_N - N\mu}{\sigma\sqrt{N}}, \tag{2}$$

tends to the standard normal as $N \rightarrow \infty$. That is,

$$P\left\{\frac{x_1 + x_2 + x_3 + \ldots + x_N - N\mu}{\sigma\sqrt{N}} \leq a\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx \quad \text{as} \quad N \rightarrow \infty. \tag{3}$$

[14] Note that in this formal definition, the sample mean multiplied by the sample length is removed from the numerator, centering the normal about the origin. The power of this result is that it holds independent of the parent distribution, provided the first and second moments are finite, and that in practice this limiting behavior can be seen even at relatively small N for many real systems. If the central limit theorem, including the assumption of independent events, holds then

$$\sigma_{sample} \approx \frac{\sigma_{parent}}{\sqrt{N}} \propto N^{\eta}\sigma_{parent}; \quad \text{with} \quad \eta = -0.5. \tag{4}$$

[15] The generalized central limit theorem describes the sum of random variables from parent distributions with power law tails, and hence infinite second moment or variance. These converge to an alpha-stable Levy distribution [*Gnedenko and Kolmogorov*, 1968]. Gaussian convergence and the regular central limit theorem are special cases of this generalized central limit theorem for finite second moment.

[16] Combining the central limit theorem and the Law of Large Numbers, the distribution of sample means in real data converges to a normal distribution centered about the mean at a rate of $1/\sqrt{N}$ as $N \rightarrow \infty$.

[17] We demonstrate this with an example which defines our null hypothesis for the convergence of earthquake interevent times. The distribution that best describes the distribution of earthquake interevent times for stationary periods, $f(\delta t)$ is the gamma distribution [*Corral*, 2004].

$$f(\delta t) = C\delta t^{\gamma-1} \exp\left(-\delta t^{\delta}/B\right). \tag{5}$$

[18] *Corral* [2004] noted that the parameters describing this distribution, when rescaled by the mean interevent time

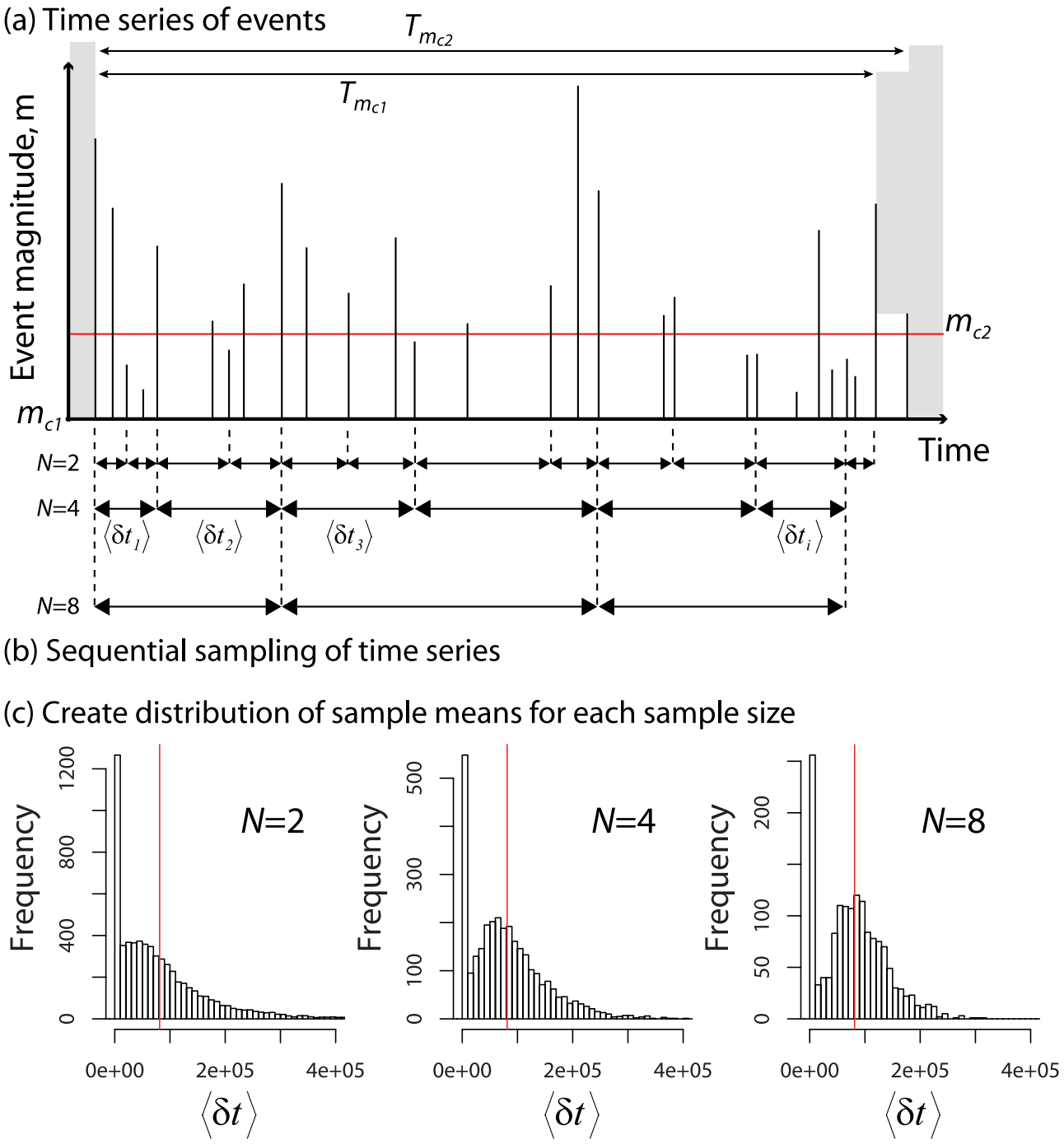

Figure 3. (color online) Illustration of how the distribution of sample means is created from a time series of earthquake events for the case of sequential sampling.

and normalized by the bin width in linear space, are the shape parameter, γ and scale parameter, B and the power $\delta \approx 1$. Using the analytic values for the mean, $\mu = \gamma B$ and variance, $\sigma = \gamma B^2$ of the gamma function, these values imply a coefficient of variation in the mean interevent times of $CV = \sigma/\mu \approx 1.2$, i.e., the range of interevent times is significantly greater than the mean interevent time. Note that this coefficient of variation is calculated for the raw data which is different to that presented in Figure 2 which was calculated for monthly averages. Even though this distribution is unlikely to be universal in the most general sense [*Hainzl et al.*, 2006] it is adequate for our purpose because it captures the first order mean, variance and skew of the data which is the information that modifies the central limit theorem convergence.

[19] Using this parent gamma distribution we randomly generated a distribution of 1000 sample means for each of the different sample lengths, N (Figure 4). We plot the sample length and standard deviation of the distribution of sample means on a log-log scale to test for power law

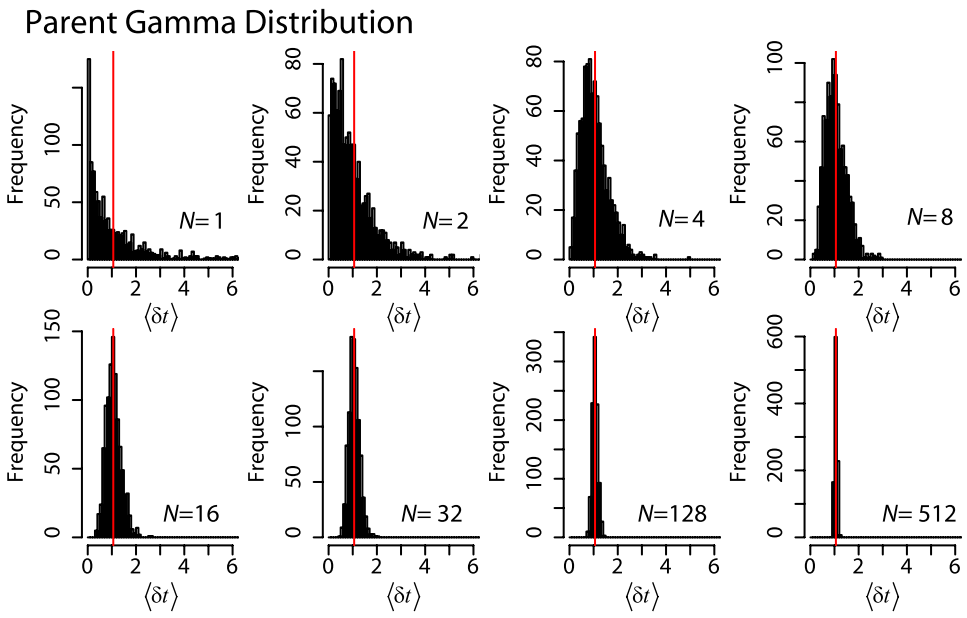

Figure 4. (color online) Demonstration of the central limit theorem using sample of means from a parent gamma distribution. The distribution of sample means tends to a normal distribution centered about the mean of the parent distribution with $\sigma_{\text{sample}} = \sigma_{\text{parent}}/\sqrt{N}$. This random sampling of the gamma distribution represents the null hypothesis for the rate of convergence of earthquake interevent times.

dependence in the rate of convergence with sample size (Figure 5a). In agreement with the central limit theorem, the distribution of sample means quickly tends to a normal distribution as $N \to \infty$ at a rate of $1/\sqrt{N}$ (i.e., $\eta \approx -0.5$ in Figure 5b). This constitutes our null hypothesis for the rate of convergence of randomly sampled interevent times from a parent gamma distribution which we compare with sequential earthquake interevent times that might be more

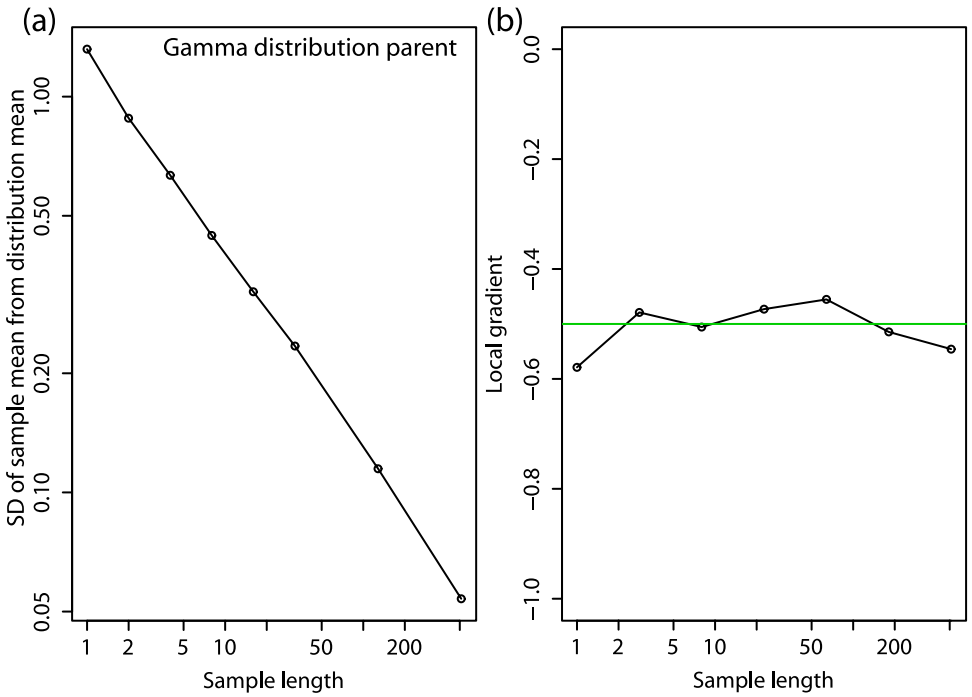

Figure 5. (color online) Demonstration of the central limit theorem prediction that the distribution sample mean converges as $\sqrt{N}$ using the sample mean distributions in Figure 4. The solid line shows the standard deviation of the sample mean distribution from the mean of the parent distribution derived from Figure 4 with varying sample length N. The gradient of the line gives the power for the rate of convergence with sample size and compares well with the central limit prediction of -0.5 (green).

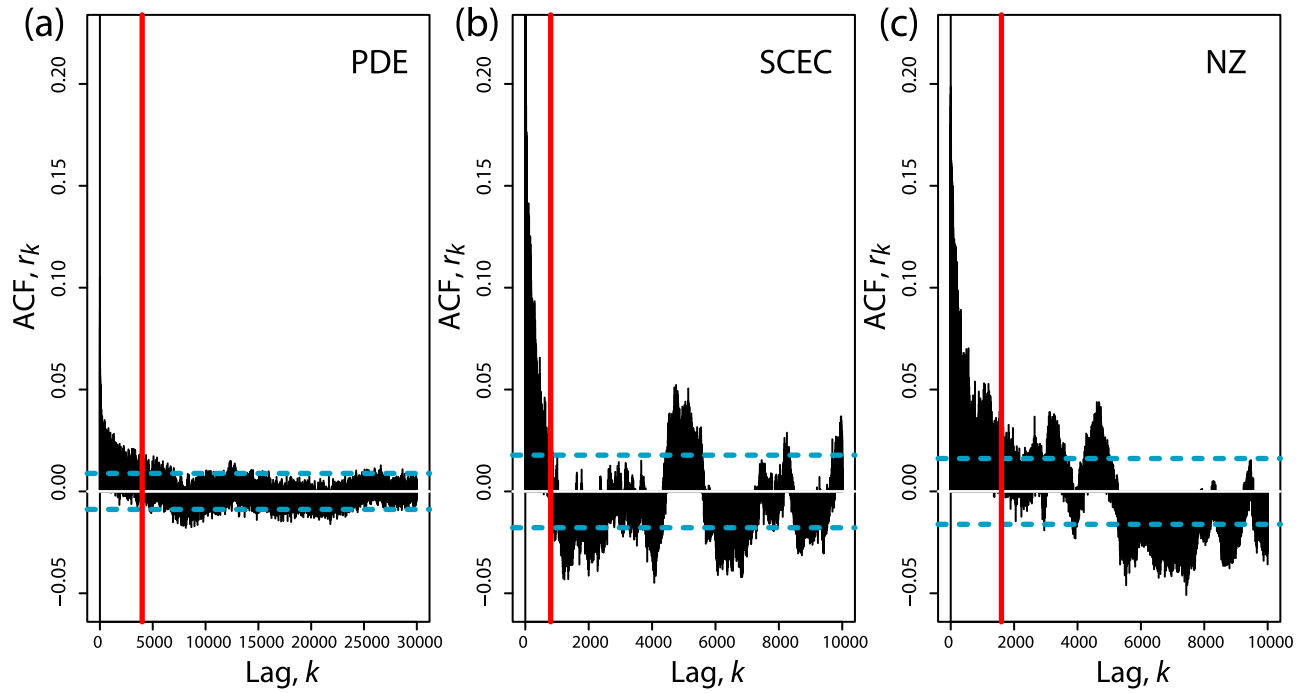

Figure 6. (color online) Autocorrelation function for the PDE, Southern California, and New Zealand catalogues. The horizontal dashed lines (blue online) show 95% confidence limits for a random process. The vertical line (red online) shows where we have chosen to have the correlations decayed; the results are insensitive to the precise choice of this point. The correlations in the PDE catalogue are weaker but longer-lasting than those in the Southern California or New Zealand catalogues.

strongly affected by temporal correlations such as aftershock sequences.

3.2. Convergence With Correlations: Autocorrelation and Effective Sample Length

[20] The autocorrelation function (ACF) provides an empirical measure of the strength and range of correlations in a data set. The strength of interactions between time series data spaced at a lag k can be quantified using the autocovariance coefficient which is defined as the average product of departures at times t and $t + k$,

$$c_k = \frac{1}{N}\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x}). \tag{6}$$

[21] The autocovariance with no lag ($k = 0$), c_0, returns the variance of the data. For nonzero lags, persistence is indicated by positive values of c_k and antipersistence is indicated by negative values of c_k. For nonzero lags, a lack of correlation is indicated by $c_k = 0$ in an infinite sample and c_k less than some limit defined by the 95% confidence intervals for finite samples in which counting errors must be taken into account [e.g., *Zivot and Wang*, 2006].

[22] The autocorrelation function, at lag k, is a variance normalized version of the autocovariance at lag k,

$$r_k = \frac{c_k}{c_0}. \tag{7}$$

[23] The expected variance for a sample of length N can be calculated for data with correlations by summing the autocorrelation function over the first N terms and multiplying by the variance in the data [*Bouchaud and Potters*, 2001]. Modifying the central limit theorem to include correlations produces

$$\sigma_N^2 = c_0\left(\frac{1}{N} + \frac{2}{N}\sum_{k=1}^{N}\left(1 - \frac{k}{N}\right)r_k\right). \tag{8}$$

[24] Note that in the absence of correlations ($r_k = 0$ for $k \neq 0$) this relation returns the central limit theorem rate of convergence. It is possible to devise sampling schemes for nonindependent variables, e.g., mixing processes, which in principle could pass the null hypothesis of Gaussian convergence. Our analysis shows that this is not the case for earthquake interevent times.

[25] Using this information about the statistical nature of the correlations we can define an effective sample length, N' which estimates how many independent pieces of data there are in the interevent time catalogue and is dependent upon the length of the sample under investigation. This effective length is defined by setting equation (8) equal to c_0/N' and rearranging. The process is non-Markovian in that future events are dependent upon more than just the current event, thus the effective sample size needs to be a function of autocorrelation coefficients over a range of lags that define the decay of the statistically significant correlations in Figures 6a–6c.

$$N' = N \Big/ \left(1 + 2\sum_{k=1}^{N}\left(1 - \frac{k}{N}\right)r_k\right). \tag{9}$$

[26] We will demonstrate that combining this effective sample size with the central limit theorem provides an appropriate correction for estimating the errors on earth-

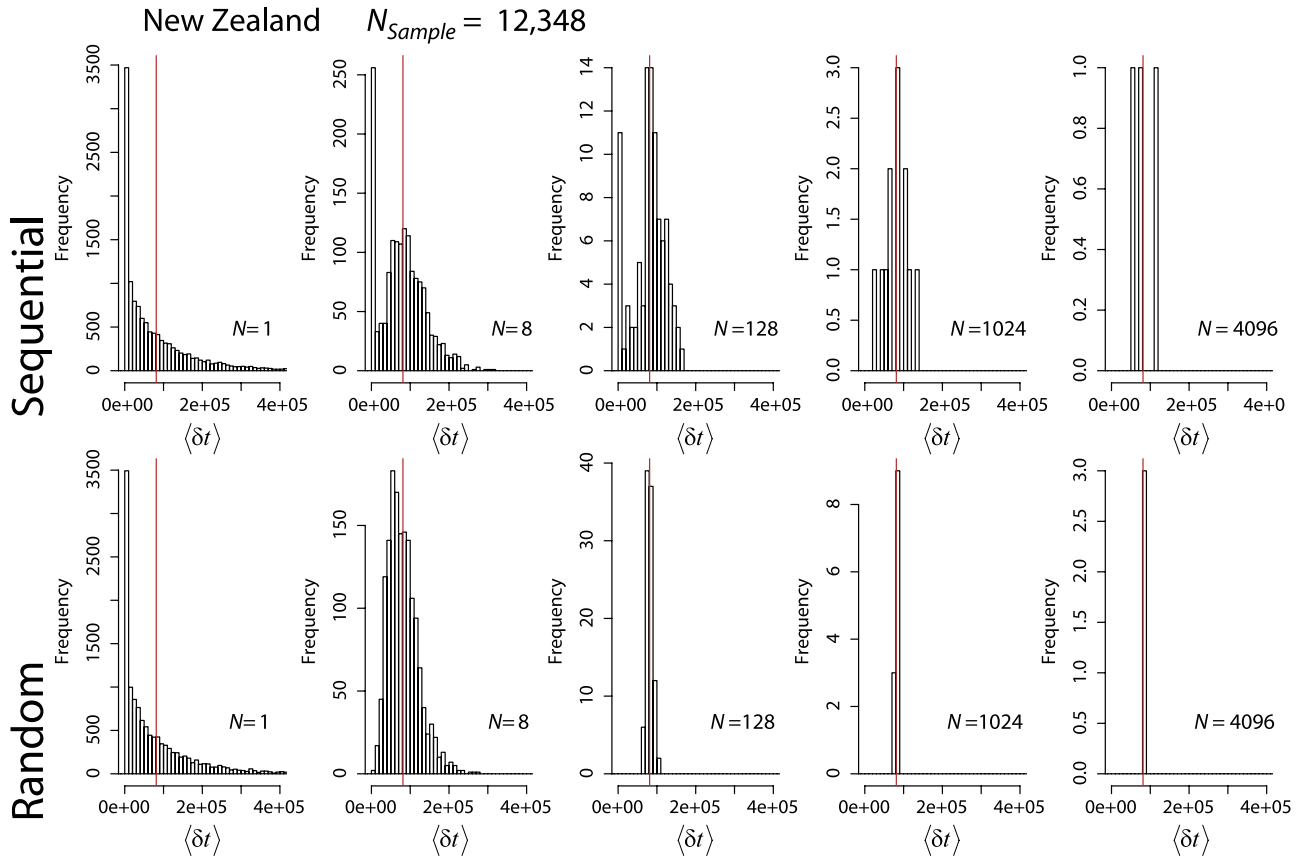

Figure 7. (color online) Examples of how the distribution of sample means varies as the sample length N tends to infinity for the (top) sequentially and (bottom) random processed data in New Zealand.

quake mean interevent times. For positive correlations $N' \leq N$, and the true error estimate is bigger than that expected from the standard deviation based on N.

4. Analysis

[27] The analysis will be carried out on three earthquake catalogue subsets (Table 1). There are two regional catalogues, New Zealand and Southern California, and one global catalogue, PDE. All of the analysis is performed using R and SSLib. We also demonstrate that the results are relatively insensitive to the precise choice of magnitude cutoff.

4.1. Sequential and Random Sampling From Earthquake Catalogues

[28] The central limit theorem assumes that the events in a sample are independent. However, aftershocks and foreshocks are known to generate correlations in the earthquake record. This has been utilized in stochastic computational models of earthquake time series, e.g., the Epidemic Type Aftershock Sequence (ETAS) model [*Ogata*, 1988], which simulate the earthquake record as a random background Poisson process from which aftershocks are stochastically produced in a branching process until no more aftershocks are produced. Such schemes imply that non-nearest neighbor events in the earthquake record may be correlated; consequently interevent times may also be correlated. In order to investigate the effect of these correlations on the convergence of mean interevent time we analyze the sequential catalogue and a randomized version in which the interevent times are randomly shuffled to remove temporal correlations.

[29] In order to compare the sequential and random interevent times across different catalogues, we must define appropriate length subsets of the main catalogues such that the subset length is a multiple of all of the sample lengths. We therefore choose to analyze subsets that are multiples of 2 from an initial sample size of 1 up to 4096, i.e., $N \in$ (1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096). Table 1 lists how many times samples of size 4096 can be drawn from each catalogue, N_{4096}. We then extract a sequential subset from the catalogue of length $N_{Sample} = 4096 \times N_{4096}$ from which we calculate random and sequential sample mean distributions. The disadvantage of this approach is that we are not using all available data in the time series. The advantage is that we ensure that we can make a fair comparison of the different catalogues in order to substantiate the conclusions of this paper and develop a technique for calculating an effective error, and predicting its convergence.

[30] For the case of the sequential catalogue, sample means are calculated for each sample length N using consecutive interevent times (Figure 3). The number of events in each distribution varies with the sample size and is given by N_{Sample}/N. Our choice of N and N_{4096} also ensures that this is an integer.

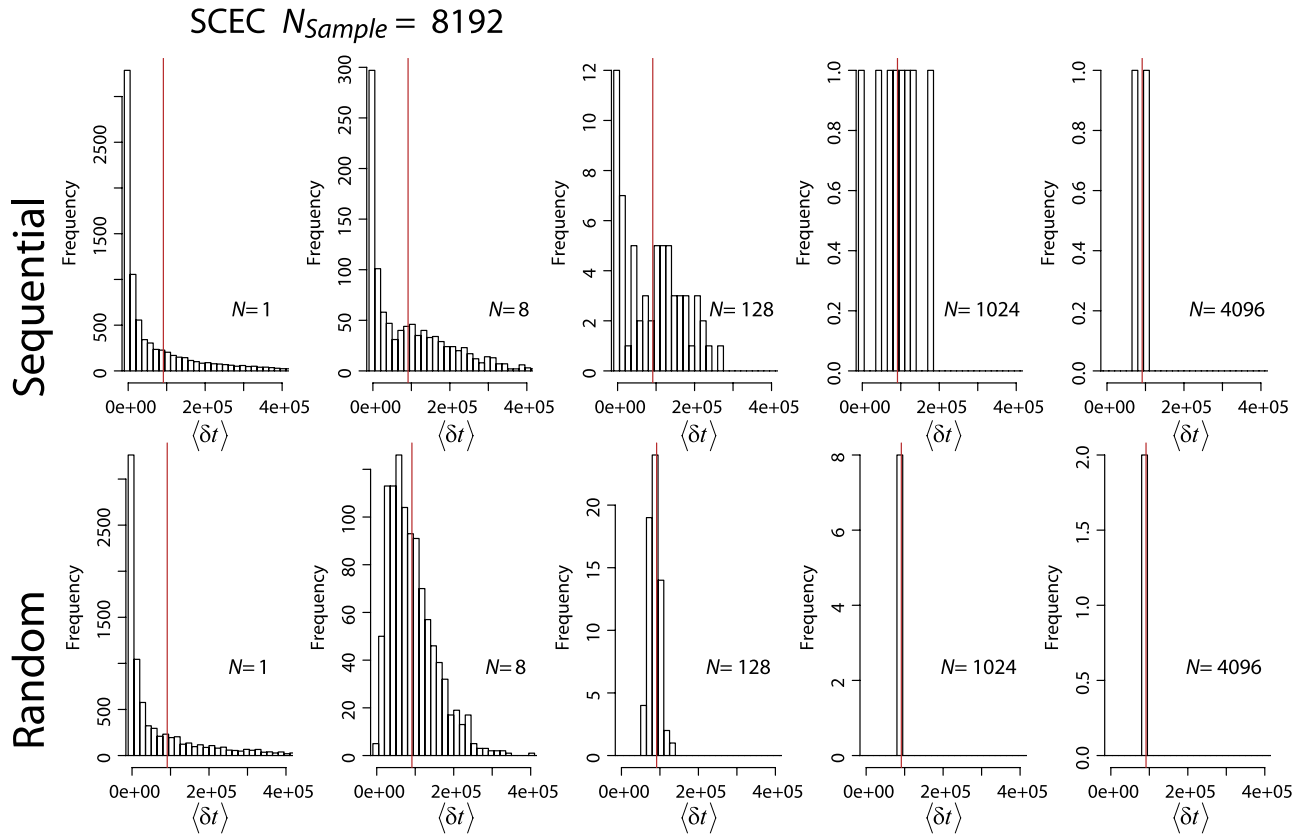

Figure 8. (color online) Examples of how the distribution of sample means varies as the sample length N tends to infinity for the (top) sequentially and (bottom) random processed data in Southern California.

[31] For the case of the random catalogue, sample means are calculated for each sample length N using the same catalogue subset as was used in the sequential analysis but the sample selection is now random without replacement, a process equivalent to randomly shuffling the catalogue. This ensures that the mean of the random and sequential distributions is the same.

[32] Typical variations in the sample mean distributions for the random and sequential cases for representative value of N are shown for the NZ, SCEC and PDE catalogues in Figures 7 and 9, respectively. The randomly processed sample means visually agree well with the predictions of the central limit theorem (compare Figures 7–9 with Figure 4). However, the sequential sample mean histograms converge more slowly than $1/\sqrt{N}$, and do not reach this limit even at $N = 4096$. This is very far from the $N > 10$ "rule of thumb" for Gaussian convergence referred to earlier.

[33] We quantify the rate of convergence to the Gaussian, as we did in Figure 5, by taking the standard deviation of the distribution of sample means about the distribution mean as a function of sample size (Figure 10a). The sequential processed rates (Figure 10a, gray lines (red online)) converge significantly slower than the randomized rates (Figure 10a, black lines). The local gradients of Figure 10a are shown in Figures 10b–10d to estimate the observed rate of convergence. The randomized data consistently fluctuates about the null hypothesis rate of $\eta = -0.5$. In contrast to the randomized case, the sequential rates of convergence are significantly slower (shallower slopes) and vary with sample size. In some, there is a tendency toward the central limit (steepening slope) for the largest sample sizes.

[34] In relative terms, the sequential PDE catalogue converges faster than either the sequential New Zealand or Southern California earthquake catalogues for a given value of N (Figure 10a) most likely because it is dominated by a greater proportion of independent events [*Huc and Main*, 2003]. However, it appears that for the largest samples the local gradient for the PDE catalogue is still rising (Figure 10b) in contrast to the NZ and SCEC catalogues where the rate has stabilized or is falling toward the value predicted by the central limit theorem (Figures 10c and 10d). Thus Gaussian convergence ($\eta \rightarrow -0.5$) need not occur at a fixed value of N for different regions.

[35] We interpret this difference between catalogues as a finite size effect that is dependent on the geographic extent of the catalogue subset being analyzed. The PDE catalogue has global coverage; therefore in the short term there are fewer correlations than the regional catalogues. This increased independence of the events in the PDE catalogue would be expected to produce a faster convergence compared to regional catalogues that contain a larger fraction of dependent events.

4.2. Dependence of Convergence Rate on Magnitude Cutoff

[36] We perform the analysis using subset catalogues with different magnitude cutoffs for the random and sequential

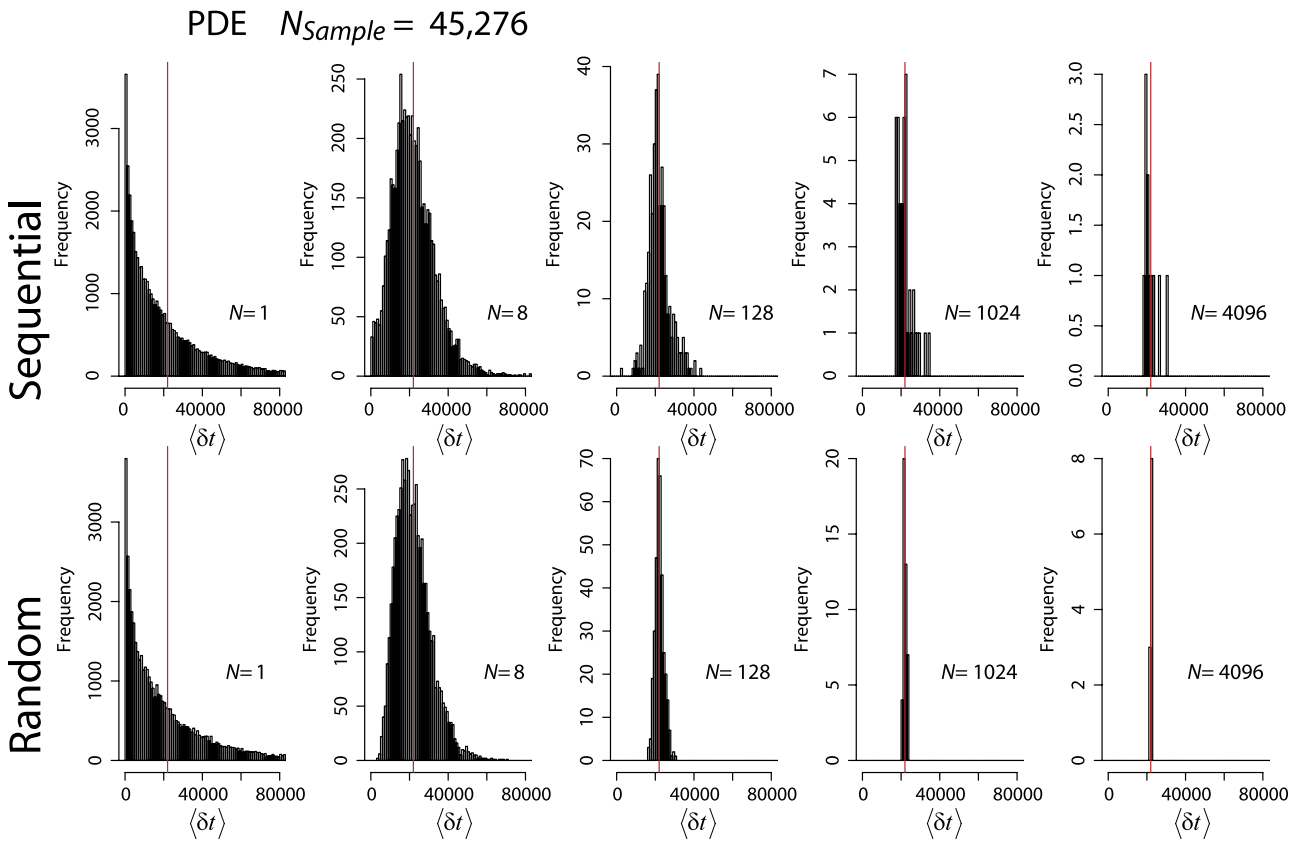

Figure 9. (color online) Examples of how the distribution of sample means varies as the sample length N tends to infinity for the (top) sequentially and (bottom) random processed data in the global PDE catalogue.

sampling. The randomly sampled models show no statistically significant variation in the rate of convergence for any of the catalogues as the magnitude cutoff is varied (Figure 11). The growing errors at longer sample lengths derive from the increase in counting errors.

[37] The sequentially sampled New Zealand and Southern California catalogue show little variation in convergence rate with magnitude cutoff (Figures 12a–12d). In contrast, the global PDE catalogue experiences faster convergence as the magnitude cutoff is raised (Figures 12e and 12f). This finite size effect arises because a significantly higher proportion of events are correlated in the regional catalogues (see sections 2 and 4.1). Where a smaller proportion of events are correlated in the PDE catalogue, increasing the magnitude threshold significantly reduces the number of correlated pairs since the larger events are more likely to be main shocks.

4.3. Investigating the Effect of Correlations Using the ETAS Model

[38] In order to test whether typical stochastic earthquake event models capture the non-Gaussian convergence of mean interevent times observed in natural catalogues, we also investigated convergence in the Epidemic Type Aftershock Sequence (ETAS) model [*Ogata*, 1988]. The ETAS model applies a conditional intensity function to seed aftershock events. The key two components of the model are a background Poisson rate, μ and the generation of aftershocks conditional on past events. The aftershock model incorporates several empirical seismological relationships:

[39] 1. Gutenberg-Richter law through the b value which describes the distribution of observed seismic moment as a power law.

[40] 2. The modified Omori's law which empirically describes the number of triggered events following a main shock as a power law in time,

$$n(t) = \frac{K}{(c+t)^p}. \tag{10}$$

where K and c are constants. Increasing p increases the rate at which aftershocks decay. This defines the parameter $A = n(t = 0) = K/c^p$.

[41] 3. The productivity law weights the triggering contribution of any event by its magnitude as $e^{\alpha m}$.

[42] These relations are combined in the conditional intensity function by summing over the history of past events, H_t at times t_i as

$$\lambda(t|H_t) = \mu + K \sum_{i:t_i<t} \exp(\alpha(m_i - m_0))(t - t_i + c)^{-p}. \tag{11}$$

m_i is the magnitude of the past event and m_0 the lower magnitude cutoff for the generated sequence.

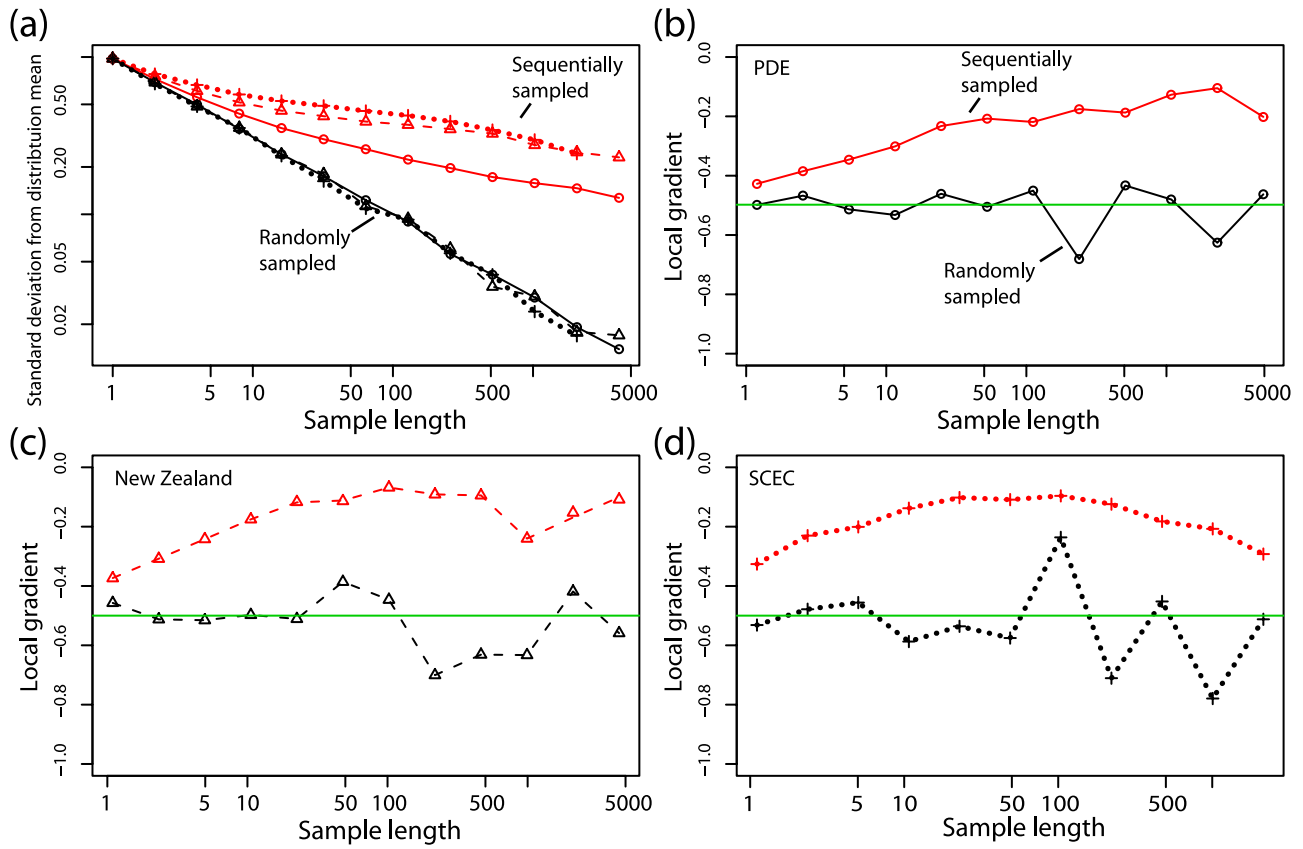

Figure 10. (color online) Standard deviation of the sample mean distribution from the mean of the parent distribution with varying sample length for (black) randomly sampled and (red) sequentially sampled earthquake data for (circle) PDE, (triangle) New Zealand, and (cross) SCEC. The randomly sampled data agree with the null hypothesis presented by the central limit theorem of convergence (see Figure 4). All of the sequential data converge slower than predicted by the central limit theorem.

[43] Stability in the event rate is governed by the branching ratio, the average number of aftershocks per event, which is given by [*Sornette and Werner*, 2005]

$$n = \frac{Ac}{p-1}\frac{\beta}{\beta-\alpha}\frac{1-\exp(-(\beta-\alpha)(m_{\max}-m_0))}{1-\exp(-\beta(m_{\max}-m_0))} \quad \text{for } p>1 \tag{12}$$

where $\beta = b\ln(10)$ and $m_{\max}$ is a maximum magnitude cutoff. If we do not impose a magnitude cutoff, we get

$$n = \frac{Ac}{p-1}\frac{\beta}{\beta-\alpha} \quad \text{for} \quad p>1, \beta>\alpha \tag{13}$$

The condition for stability is that n should be finite and less than 1. Firstly, to have a finite n we require $p > 1$ if $\beta > \alpha$ (in the absence of a magnitude cutoff); secondly, the parameters A, c, p, α and β must have values that combine to give $n < 1$ [*Sornette and Helmstetter*, 2002].

4.3.1. Extrapolating the Record

[44] We generated a set of 5 ensemble long synthetic catalogues using the ETAS model to investigate how earthquake interevent times converge for longer records than are available in current earthquake catalogues, where the parameterization is chosen such that the interevent time distribution is known to be stationary over the time frame investigated. Analysis was only performed on data after initial transients associated with the model run in phase had decayed. Since we are interested in the generic behavior of earthquake like interevent correlations we choose a representative parameterization of the ETAS model, rather than trying to simulate a specific catalogue, with $b = 1$, $\mu = 0.5$, $A = 10$, $\alpha = 0.9$, $c = 0.01$, $p = 1.2$. These values represent "typical" ones in the middle of the range of reported values in the literature, and correspond to a branching ratio of $n = 0.821$, thus ETAS model is stable. The synthetic catalogues generated here correspond to a 430 year simulation length. On average these contain 363747 events, much more than in any real catalogue. These events range over 6 orders of magnitude above an arbitrary magnitude threshold. Since this record is substantially longer, we now investigate the behavior up to and including sample means containing $N = 32768$ events.

[45] The qualitative variation of the histograms produced using the randomly and sequentially sampled synthetic ETAS time series (e.g., Figure 13) are comparable to that for the real earthquake catalogues analyzed (e.g., Figures 7 and 9).

[46] The rates of convergence of the sample mean are shown in Figures 14a and 14b for five realizations of the ETAS model. The ensemble runs clearly demonstrate the

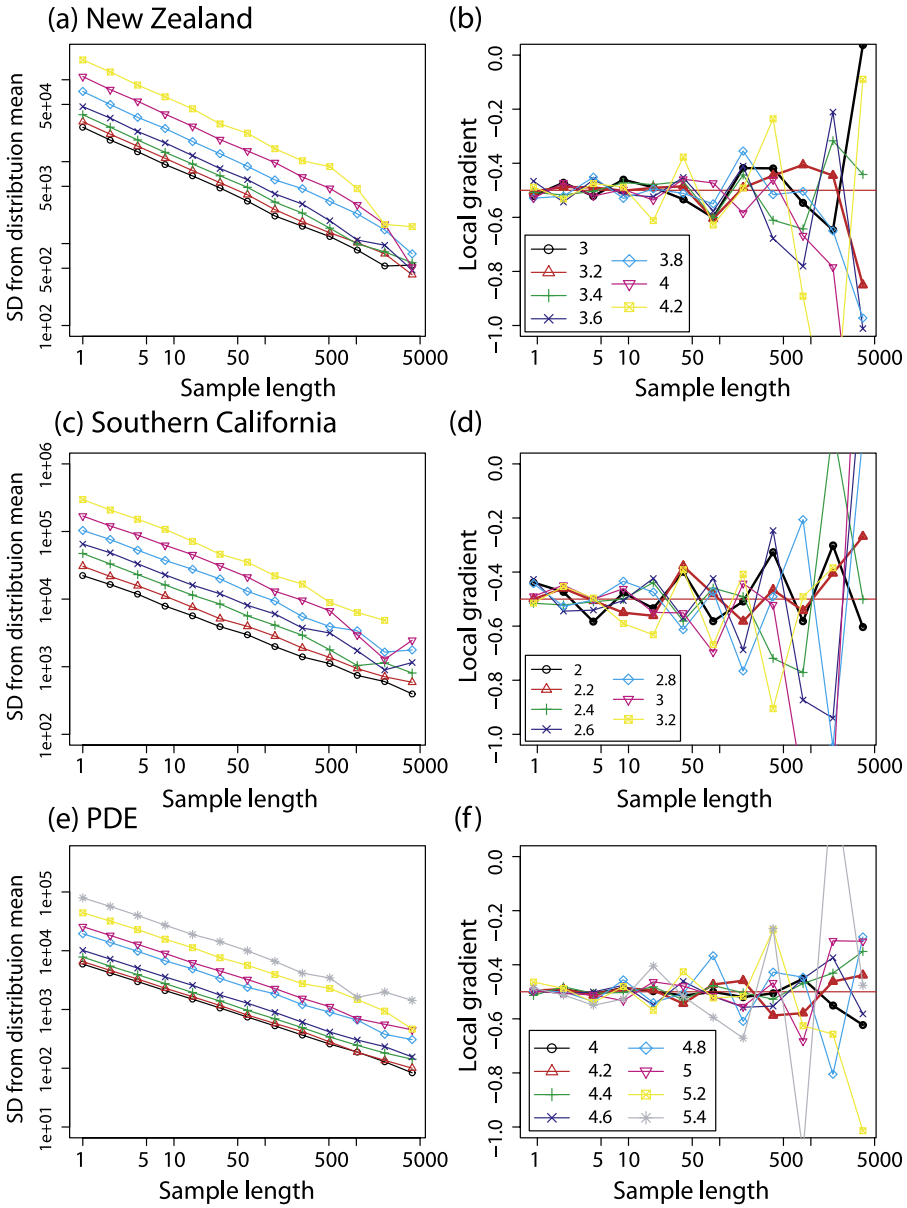

Figure 11. (color online) Rates of convergence of mean interevent times for randomized catalogues with varying magnitude cutoff.

effects of growing counting errors as the sample size increases, making it hard to discern the actual rate of convergence for the largest sample sizes. The ensemble mean of these runs is added to mitigate the effects of counting errors. Once again, the random model agrees well with the central limit theorem prediction of convergence at a rate of $1/\sqrt{N}$, i.e., the local gradient fluctuates about $\eta = -0.5$ (Figure 14b). The rates of convergence are slower for the sequential sampling.

4.3.2. Effect of Varying Strength of Correlations

[47] The ETAS model also allows us to investigate the effect of varying the strength of correlations in the interevent time series.

[48] Increasing p increases the rate at which aftershocks decay and hence decreases the ratio of dependent to independent events. Thus decreasing p slows the rate of convergence (Figures 14c and 14d). The longer lived correlations associated with lower p increase the duration of the non-Gaussian convergence as longer event records become necessary in order to increase the proportion of uncorrelated to correlated events to the same degree.

[49] Further, in the model the number of aftershocks produced by an event depends on the magnitude of that event as $e^{\alpha m}$. For higher values of α, there will be a larger number of correlated events (aftershocks) following an event of a given magnitude, and the temporal duration of the sequence is therefore longer assuming the decay rate p is kept constant. This effect increases the average correlation length, but also makes it more variable with event magnitude.

[50] Figures 14c–14f referred to above all show ensemble runs for each of the model parameterizations. Fluctua-

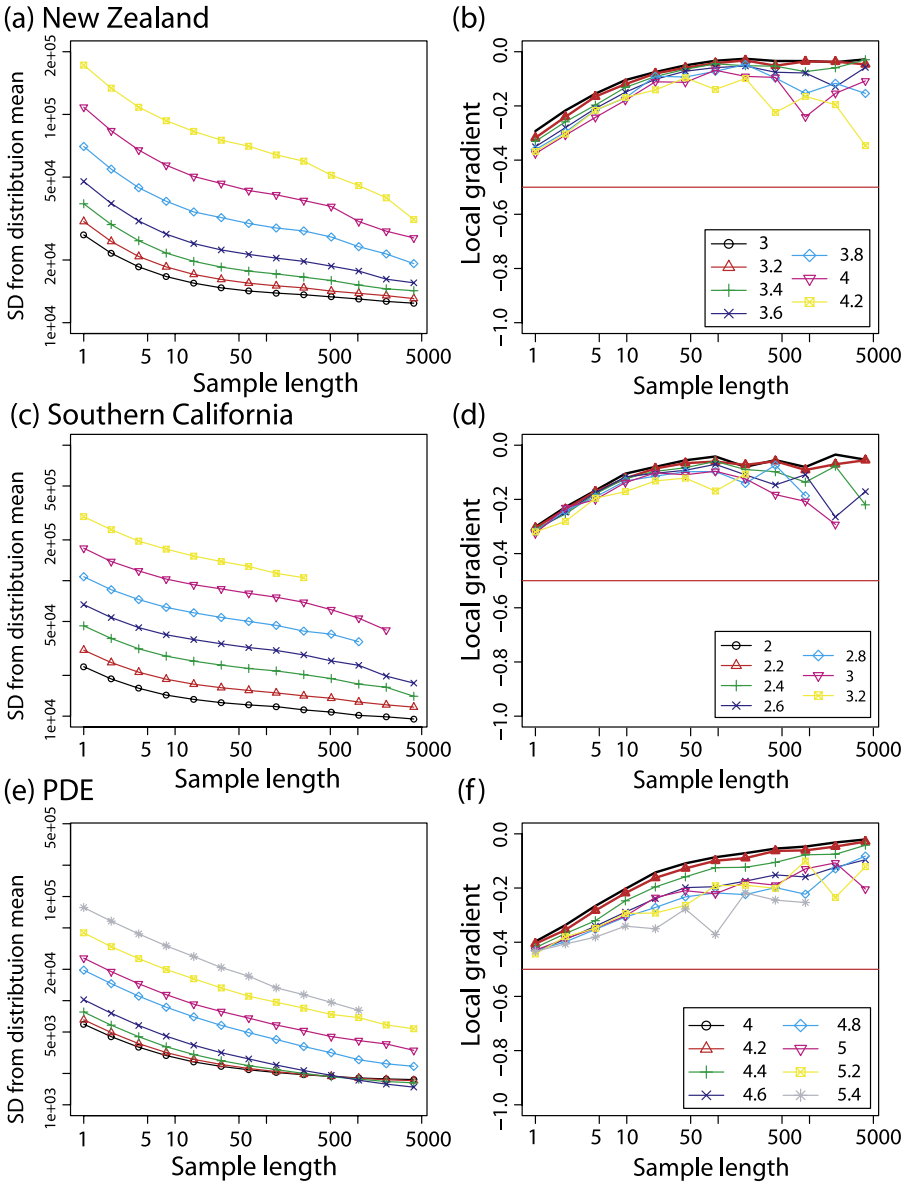

Figure 12. (color online) Rates of convergence of mean interevent times for sequentially sampled catalogues with varying magnitude cutoff.

tions between the ensemble runs at a given p or α are less than the differences as p or α are varied, as shown in Figures 14c and 14e. Thus the convergence rate and evolution for a given parameterization is robust against statistical variability in the ETAS model and provides a technique for validating regional studies that have been fitted to the ETAS model.

4.3.3. Summary of ETAS Model Convergence

[51] Overall, the ETAS model does a good job in reproducing the convergence observed in real data. It converges at a slower rate than the random case due to the presence of many correlations in the shorter samples. As the sample size increases, the ensemble rate of convergence increases toward $1/\sqrt{N}$. Such convergence depends critically on the sample period being sufficiently long compared to the average time between extreme events that generate many aftershocks, and may not occur at a particular universal value of N. However, simulation using the ETAS model is not very good for predicting errors on the mean interevent times for large sample sizes because of computational constraints on running enough large-scale simulations to reduce counting errors to a suitable size (see growing fluctuations with sample size for sequential sampling in Figure 14b, dashed lines (blue online)).

5. Error Prediction Using Autocorrelation Function

[52] The correlogram (plot of the autocorrelation function, r_k defined in equation (7) as a function of lag k) for each of the catalogues is plotted in Figure 6 to quantify the strength and range of interevent time correlations. In the

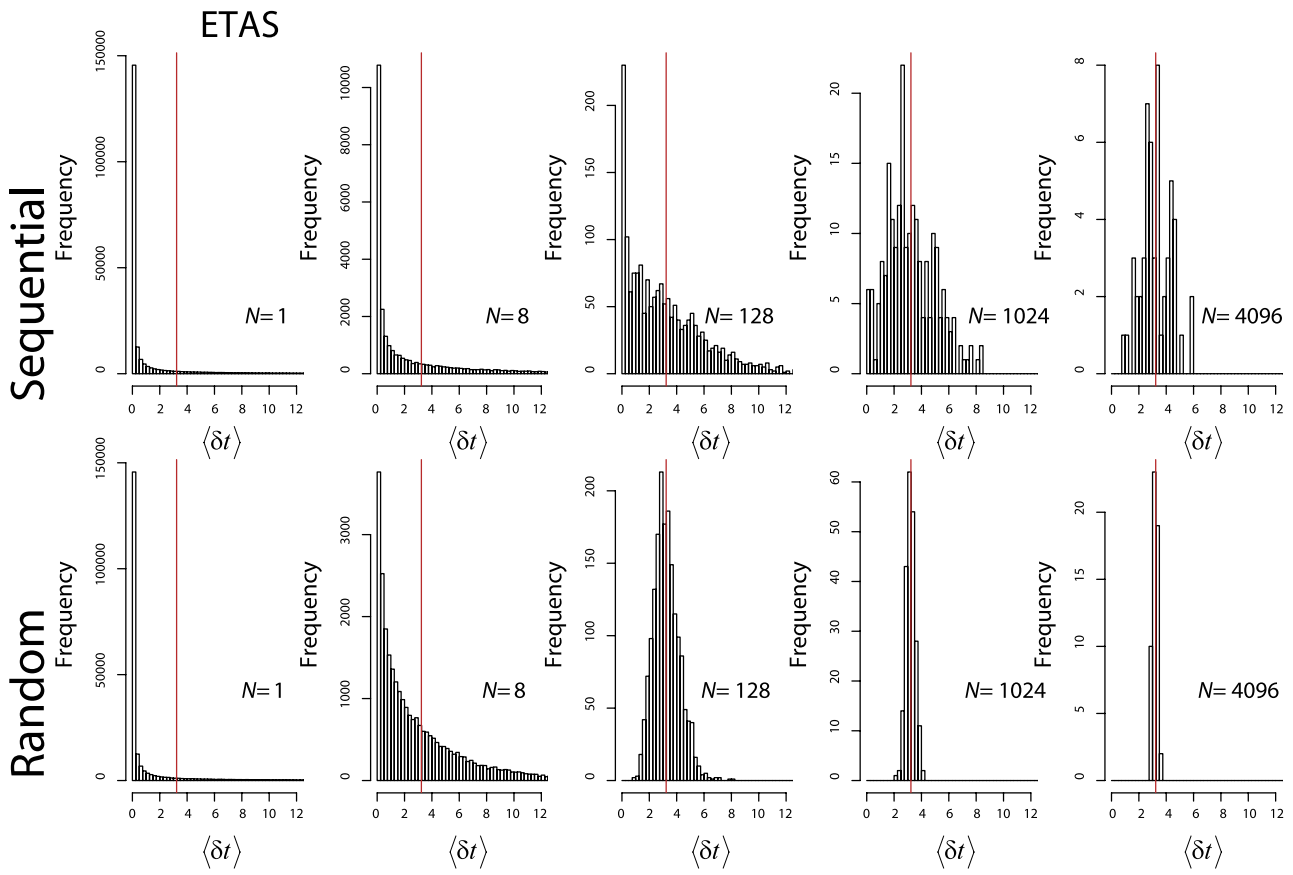

Figure 13. (color online) Examples of how the distribution of sample means varies as sample length N tends to infinity for (bottom) random and (top) sequentially data derived from the ETAS model.

time domain, r_k identifies that the global PDE catalogue contains weaker but longer range correlations than the NZ and SCEC catalogues (Figure 6). Thus r_k can be used quantitatively to investigate temporal finite size effects. The horizontal dashed lines (blue online) show the 95% confidence limits about 0 based on the assumption that the data is random and uncorrelated [e.g., *Zivot and Wang*, 2006], thus decaying correlations above the dashed line (blue online) at short lag times can be accepted with this level of confidence. Below these confidence limits the correlations are not statistically significant. As the lag is increased counting errors in the autocorrelation coefficients grow, generating spurious fluctuations in the moving average, so we must also crop the useful data, at a lag of length k_c, where the correlations first decay below the confidence intervals. The position of the cropping is slightly arbitrary, due to high frequency fluctuations in r_k, and is marked on Figure 6 as the vertical line (red online). The longer range, almost periodic, moving average diversions above the confidence limits to the right of the cropping point (e.g., Figure 6b shows several clear excursions above the confidence limits) are artifacts arising from a finite sample size. The autocorrelation function should be used to verify that a correlation length measured from data is real and not just one of these sampling effects.

[53] For the catalogues investigated here, summing the autocorrelation function over the significant range of non-zero lags (i.e., summing r_k from $k = 1$, k_c) indicates that each event is on average correlated with 74.0 events for New Zealand over a range of ∼1600 events, 48.6 events for southern California over a range of ∼800 events and 49.8 events for PDE over a range of ∼4000 events. It is important to stress that the results are insensitive to the precise choice of the position of this cutoff. This defines the range of the correlations to be on the order of a few thousand events.

[54] For approximately Gaussian distributed data with correlations, the effective sample length N' defined by equation (9) can be used to define convergence in the Central Limit form as $1/\sqrt{N'}$, which is plotted in Figures 15a and 15b. The standard deviation predicted by the effective sample length compares well with that for the real data (see Figures 10, 15b, and 15c). Errors are largest and deviate most from $1/\sqrt{N}$ convergence for small sample sizes in regional catalogues. The errors for all catalogues tend to converge for larger sample sizes. Thus we have demonstrated that we can reproduce the observed rate of convergence using only knowledge of the variance and autocorrelation function defined by the data.

[55] This use of the autocorrelation function to generate an effective sample size as a function of sample length reproduces the observed errors well and allows us to predict how the errors decay at larger sample sizes than are currently available. Further, this analysis motivates a com-

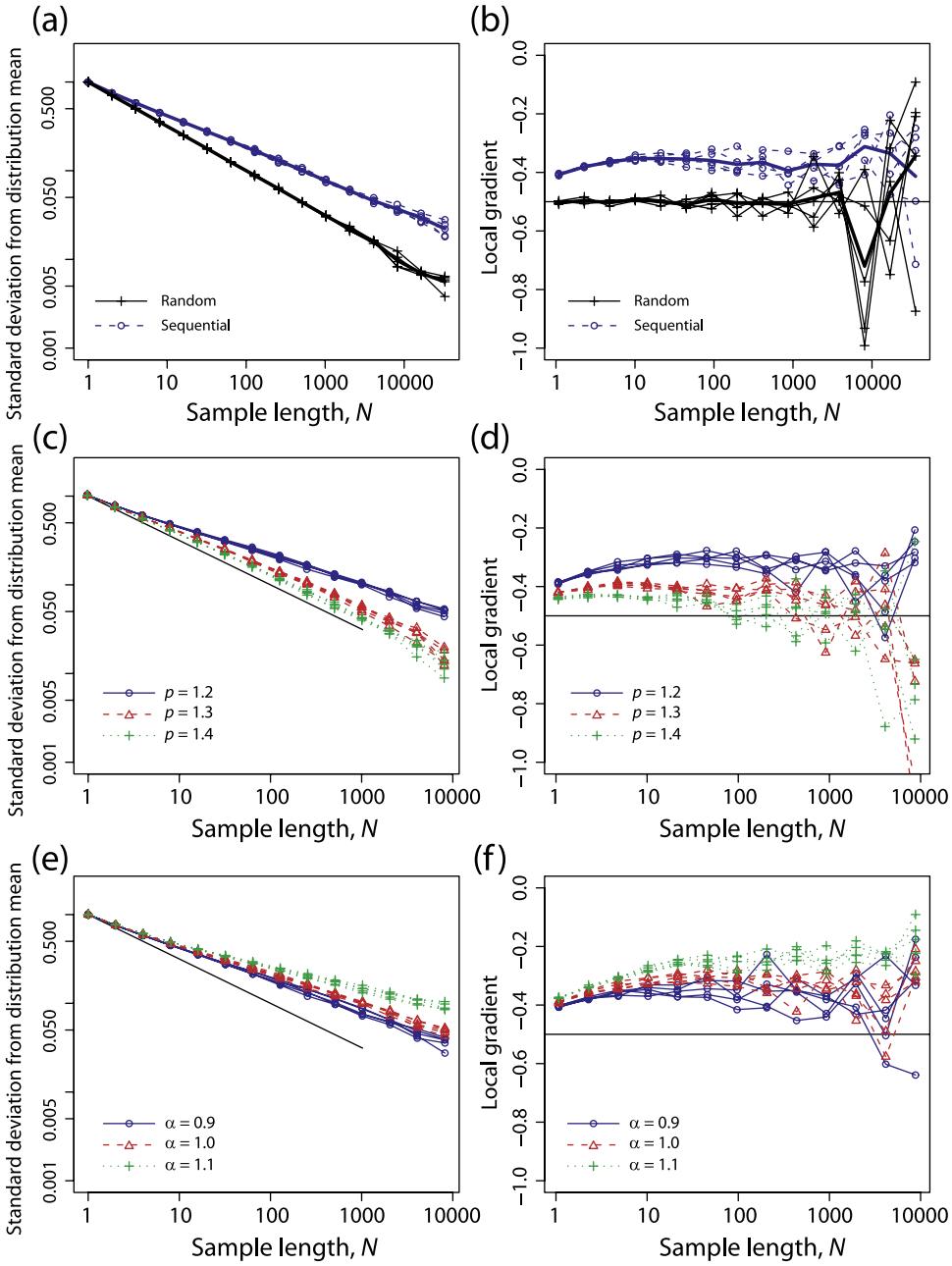

Figure 14. (color online) (a and b) Rate of convergence of sample mean using five ensemble runs of the ETAS model with the same parameters. The ensemble mean has been added for each set of runs as a thicker line. (c and d) The rate of convergence varies as the Omori decay rate is increased. (e and f) The rate of convergence varies as the dependence of aftershocks on magnitude is increased.

plementary definition of stationarity, i.e., where the autocorrelation function is temporally stable.

6. Discussion and Potential Applications

[56] The arithmetic mean interevent time, proportional to the inverse of the harmonic mean event rate, provides the most stable measure of earthquake activity and converges smoothly. This should be preferred as a measure of earthquake activity over the arithmetic mean event rate which does not converge cleanly due to its sensitivity to extreme events. Skewness in the interevent time parent distribution makes it more likely that the mean interevent time will converge from above (or the event rate from below). Thus on average we will underestimate earthquake activity from catalogue samples which have not yet converged. This could be a significant systematic error in many current estimations of time-independent seismic hazard.

[57] The underestimate of random errors on the mean are largest for small sample sizes. Time-dependent hazard assessments typically use small subsets of a catalogue, and so are very vulnerable to this source of statistical noise. Our results suggest time-dependent studies need to demonstrate that observed "anomalous" trends in activity lie

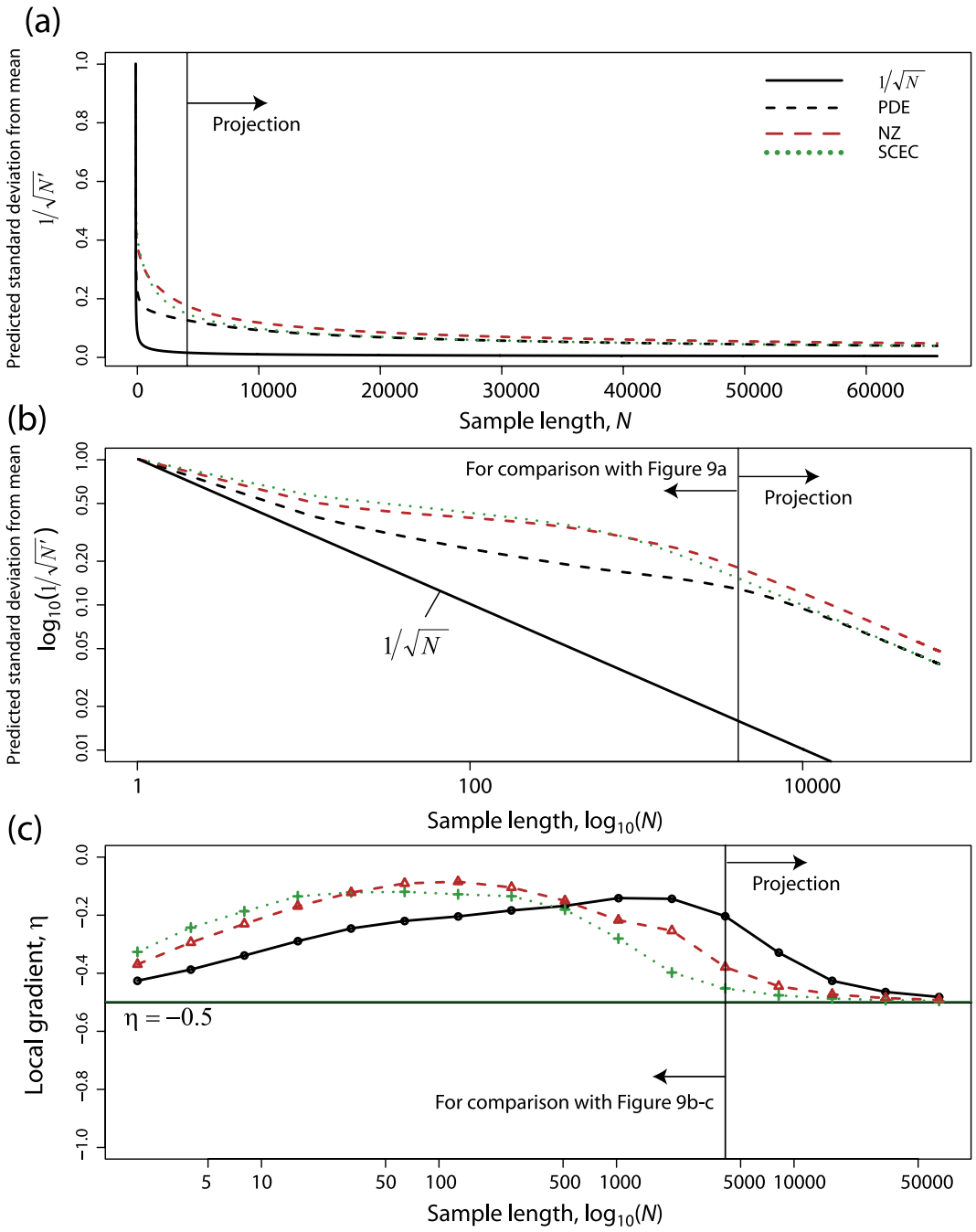

Figure 15. (color online) (a) Plot comparing $1/\sqrt{N}$ central limit convergence in the absence of correlations with $1/\sqrt{N'}$ autocorrelation corrected convergence for the PDE, NZ, and SCEC catalogues. To convert these normalized values to actual standard deviations in the mean requires that the plotted values be multiplied by the standard deviation of the underlying data. The presence of persistence in the interevent time series generates slower convergence. (b) The same as for Figure 15a but plotted on a log-log scale. The vertical black line shows the transition between data that can be directly compared with Figure 9a and the projection of the effect of the correlations on larger sample sizes. (c) The local gradients of Figure 15b for direct comparison with Figures 9b and 9c. The horizontal line at $\eta = -0.5$ shows the theoretical rate for central limit convergence in the absence of correlations.

outwith the error bounds expected from sampling a correlated time series. This sampling effect presents an additional source of difficulty in posing statistical models involving aftershock sequences as a null hypothesis to be rejected in evaluating earthquake forecasting power at a higher level, for example, from candidate "precursors" (http://www.nature.com/nature/debates/earthquake/equake_frameset.html).

[58] The autocorrelation function shows that geographically smaller, regional catalogues, such as New Zealand and Southern California, have stronger correlations over a shorter range of lags than the global PDE catalogue which

has weaker but longer lasting correlations. Generally, this will translate into larger errors for more localized geographic studies. This result has important implications for high spatial resolution seismic hazard studies, for example those currently being developed for application in catastrophe (''cat'') modeling in the earthquake reinsurance market [*Grossi et al.*, 2005]. Fortunately equation (9) can be used to estimate the proportion of correlated events and the effective error in such small samples. It can also be used a posteriori to validate the hypothesis of stationarity in time-independent seismic hazard, i.e., the autocorrelation function is not varying significantly in time.

[59] Many studies use declustering algorithms to remove aftershock events from earthquake catalogues. The technique presented here could be used as an a posteriori check on the quality of the declustering, with ideal declustering defined by a filtered catalogue (of events identified as independent by the algorithm) exhibiting Gaussian convergence, with slower rates η defining a less effective declustering algorithm. Finally the stability of the analysis with respect to varying magnitude cutoff implies a degree of self similarity with respect to the strength and range of correlations in the interevent time series, and is an indication of the robustness of the technique.

7. Conclusions

[60] Hazard assessments should try to quantify errors, but rarely do explicitly. We have shown that the rate of convergence of the mean earthquake interevent time is slower than the central limit theorem prediction of $1/\sqrt{N}$, primarily due to the presence of correlations in the interevent time series. As a consequence many studies of seismic hazard (time independent or time dependent) currently underestimate the true uncertainty in mean and standard deviation of parameters such as event rate or interevent time.

[61] We have presented a simple technique to quantify errors on earthquake mean interevent times using only the variance of the data and the autocorrelation function of the interevent time series. Specifically, the autocorrelation coefficients, r_k can be used to define an effective sample size which corrects the sample size for the number of events that it is likely to be correlated to, which changes as a function of sample size, such that $N' = N/(1 + 2\sum_{k=1}^{N}(1 - \frac{k}{N})r_k)$. This effective sample size can be used to estimate the rate of convergence of the earthquake interevent times as a function of sample size as $1/\sqrt{N'}$.

Notation

Statistical variables

$x_1, x_2, \ldots$	A sequence of random variables
N	Number of random variables/events in sample
$\mu, \bar{x}$	Mean of a sample
σ	Variance of a sample
CV	Coefficient of variation
η	Local rate of convergence

Earthquake analysis specific variables

δt	An interevent time
$\langle \delta t \rangle$	A sample mean interevent time
N_{events}	Number of events in an earthquake catalogue subset
N_{4096}	Whole number of samples of length 4096 that fit into catalogue subset

Gamma function properties

$f(\delta t) = C\delta t^{\gamma -} \exp(-{}^{\delta t^{\gamma}}/_{B})$	The generalized gamma function
γ	Shape parameter
B	Scale parameter
C	Normalization factor

Autocorrelation and effective sample size

k	Lag between events
c_k	Autocovariance at lag k
r_k	Autocorrelation at lag k
N'	Effective sample size estimating the number of independent pieces of data in a sample

ETAS parameters

b	Gutenberg-Richter b value
μ	Background rate
α	The exponent relating the production of aftershocks as a function of magnitude
p	Omori's law power, describing the decay rate for aftershocks
$A = n(t = 0) = K/c^p$	Occurrence rate of earthquakes in the Omori's law at zero lag from an event in equation (10)

[62] **Acknowledgments.** M.N. was funded by the EPSRC NANIA grant GR/T11753/01 and EU ''TRIGS'' training network. S.T. was funded by EPSRC DTA. The analysis in this project was carried out using R [*R Development Core Team*, 2007] and SSLib [*Harte*, 2007]. We would like to thank two anonymous reviewers and Rodolfo Console for constructive comments.

References

Bouchaud, J.-P., and M. Potters (2001), *Theory of Financial Risks: From Statistical Physics to Risk Management*, Cambridge Univ. Press, New York.

Corral, Á. (2004), Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes, *Phys. Rev. Lett.*, *92*, 108501.

DeMets, C. (1993), Earthquake slip vectors and estimates of present-day plate motions, *J. Geophys. Res.*, *98*, 6703–6714.

Gnedenko, V. V., and A. N. Kolmogorov (1968), *Limit Distributions of Sums of Independent Random Variables*, Addison-Wesley, Boston, Mass.

Grossi, P., et al. (2005), *Catastrophe Modeling: A New Approach to Managing Risk*, Springer, New York.

Hainzl, S., et al. (2006), Estimating background activity based on interevent-time distribution, *Bull. Seismol. Soc. Am.*, *96*, 313–320.

Harte, D. (2007), *Users Guide for the Statistical Seismology Library*, Statistics Research Associates, Wellington, N. Z.

Huc, M., and I. G. Main (2003), Anomalous stress diffusion in earthquake triggering: Correlation length, time dependence, and directionality, *J. Geophys. Res.*, *108*(B7), 2324, doi:10.1029/2000JC000237.

Laplace, P. S. (1812), *Théorie Analytique des Probabilitiés*, Courcier, Paris, France.

Main, I. G., L. Li, J. McCloskey, and M. Naylor (2008), Effect of the Sumatran mega-earthquake on the global magnitude cut-off and event rate, *Nat. Geosci.*, *1*, 142, doi:10.1038/ngeo141.

Naylor, M., and H. D. Sinclair (2007), Punctuated thrust deformation in the context of doubly vergent thrust wedges: Implications for the localization of uplift and exhumation, *Geology*, *35*(6), 559–562, doi:10.1130/G23448A.1.

Ogata, Y. (1988), Statistical-models for earthquake occurrences and residual analysis for point-processes, *J. Am. Stat. Assoc.*, *83*, 9–27.

Poisson, S. D. (1837), *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile, Précedées des Règles Générales du Calcul des Probabilités*, Bachelier, Paris, France.

Ross, S. M. (2003), *Introduction to Probability Models*, 8th ed., Elsevier, New York.

Sornette, D., and A. Helmstetter (2002), Occurrence of finite-time singularities in epidemic models of rupture, earthquakes, and starquakes, *Phys. Rev. Lett.*, *89*(15), 158501.

Sornette, D., and M. J. Werner (2005), Constraints on the size of the smallest triggering earthquake from the epidemic-type aftershock sequence model, Bath's law, and observed aftershock sequences, *J. Geophys. Res.*, *110*, B08304, doi:10.1029/2004JB003535.

R Development Core Team (2007), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, available at http://www.R-project.org.

Zivot, E., and J. Wang (2006), *Modelling Financial Time Series With S-Plus*, 2nd ed., Springer, New York.

I. G. Main, M. Naylor, and S. Touati, School of GeoSciences, University of Edinburgh, Grant Institute, West Mains Road, Edinburgh EH9 3JW, UK. (mark.naylor@ed.ac.uk)